North East Linguistics Society

Volume 31
Issue 1 *NELS 31: Volume 1*

Article 5

2001

An MEG Study of Tone Processing Asymmetries in English versus Mandarin Speakers

Benjamin Bruening
Massachusetts Institute of Technology

Elissa Flagg
Massachusetts Institute of Technology

Vivian Lin
Massachusetts Institute of Technology

Follow this and additional works at: https://scholarworks.umass.edu/nels

Part of the Linguistics Commons

Recommended Citation

Bruening, Benjamin; Flagg, Elissa; and Lin, Vivian (2001) "An MEG Study of Tone Processing Asymmetries in English versus Mandarin Speakers," *North East Linguistics Society*: Vol. 31 : Iss. 1 , Article 5.
Available at: https://scholarworks.umass.edu/nels/vol31/iss1/5

This Article is brought to you for free and open access by the Graduate Linguistics Students Association (GLSA) at ScholarWorks@UMass Amherst. It has been accepted for inclusion in North East Linguistics Society by an authorized editor of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

An MEG Study of Tone Processing Asymmetries in English versus Mandarin Speakers

Benjamin Bruening, Elissa Flagg, Vivian Lin

Massachusetts Institute of Technology[1]

1. Introduction

This study examines what effect language background may have upon early auditory processing, looking in particular at hemispheric asymmetries during the processing of fundamental frequency (F0) information by English versus Mandarin Chinese speakers. Within the psycholinguistic literature, two main approaches to the asymmetry issue may be identified: the functional approach, under which the functional context of an item determines where it is processed in the brain, and the *acoustic* approach, under which the acoustic properties of a stimulus determine where it is processed in the brain. Much of the existing research has focussed on the processing of the F0 information carried on linguistic tones, since it serves different functions in non-tone languages (such as English) versus tone languages (such as Chinese). The functional approach predicts that English and Chinese speakers will vary in their response to F0 information. The Chinese speakers are expected to process tones like all linguistic information, in the left hemisphere (LH), in contrast to the English speakers. On the other hand, the acoustic approach predicts that English and Chinese speakers will not differ in the processing of F0 information. On this view, each hemisphere is hypothesized to be computationally biased for different types of information: the LH for high-frequency information, the right hemisphere (RH) for low-frequency information; thus, F0 information is predicted to be processed in the RH for both English and Chinese speakers.

Section 2 discusses in more detail the existing evidence in favor of the functional and the acoustic theories of auditory processing. Section 3 provides background on the

[1] This project also funded by the Mind Articulation Project, International Cooperative Research Project, Japan Science and Technology Corporation, Tokyo 113-0034. E. Flagg also funded by the Social Sciences and Humanities Research Council of Canada Fellowship #752-98-0362.

© 2001 by Benjamin Bruening, Elissa Flagg, and Vivian Lin
NELS 31

tool we employ to test these theories, namely, the M100, an automatic brain response to auditory stimuli measured by magnetoencephalography (MEG). Section 4 describes the materials and methods used in our study. The experimental results, as well as our interpretation of them, are presented in Section 5, and we end with concluding remarks in Section 6.

2. Functional versus Acoustic Approaches to Auditory Processing

Early psycholinguistic research addressing the question of whether there are hemispheric asymmetries in linguistic tone processing (for speakers of tone languages) seems to converge on the conclusion that linguistic tone processing takes place in the left hemisphere. In these early studies, evidence for left-hemisphere lateralization comes from two main sources: dichotic listening experiments and subjects with unilateral brain damage. For example, Van Lancker and Fromkin (1973, 1978) show that speakers of a tone language, Thai, have a consistent right ear advantage in discrimination tasks involving minimal pairs which differ in either tonal or segmental (consonantal) content, but no particular ear advantage in discrimination tasks involving non-linguistic tones. The fact that speakers of a non-tone language, English, only exhibit a right ear advantage in tasks involving segmental contrasts, and not for tone-based minimal pairs (either linguistic or non-linguistic), led Van Lancker and Fromkin to conclude that pitch discrimination is lateralized to the left hemisphere when the pitch differences are linguistically processed (Van Lancker and Fromkin 1973). Van Lancker and Fromkin argue that the lateralization of pitch processing is based upon functional context, rather than on particular acoustic features. Since pitch information in linguistic stimuli is lexically contrastive in tone languages, pitch processing will be lateralized to the left hemisphere in tone-language speakers. In non-tone-language speakers, pitch information is not lexically contrastive (setting aside questions regarding stress placement), and thus, pitch processing will not be lateralized to the left hemisphere. If lateralization were dependent upon acoustic features, rather than functional context, we would expect Thai speakers and English speakers to exhibit similar patterns of lateralization with respect to the same pitch stimuli.

The functional theory of hemispheric specialization gains support from studies involving patients with unilateral brain damage. Comparing Thai aphasics with unilateral left-hemisphere damage to a normal Thai speaker and a Thai speaker with right-hemisphere brain damage, Gandour and Dardarananda (1983) found that the left-damaged subjects performed significantly worse than both the normal and the right-damaged Thai subjects on a linguistic tone identification task. Along similar lines, Packard (1986) found that tone processing was impaired in left-damaged Chinese aphasics, who produced the same number of errors involving tones as consonants in a word repetition task.

More recently, Gandour, Wong, and Hutchins (1998) have used positron emission tomography (PET) technology to test for differential tone processing by English versus Thai speakers. GWH ran both English and Thai subjects on three different conditions: a baseline condition which involved silence, a tone condition which involved a same-different judgment task over Thai words , and a pitch condition which involved a same-different judgment task over non-linguistic stimuli (specifically, low-pass filtered versions of the Thai words used in the experiment). For Thai speakers, the tone-minus-pitch and the tone-minus-baseline subtractions both revealed activation in the left frontal operculum (roughly Broca's area), whereas for the English speakers, the tone-minus-pitch subtraction showed no frontal activation. GWH interpret these results as support for the functional theory of hemispheric specialization: in response to tone information, the Thai speakers show LH activation which is not present for English speakers. (See also Gandour et al. 2000, which extends the work of Gandour, Wong, and Hutchins (1998) to include speakers of Mandarin Chinese as experimental subjects.)

However, all of these reported studies suffer from a serious flaw: due to limitations imposed by their various experimental paradigms, they fail to distinguish between tone processing and higher linguistic processing, for instance lexical access. All of the linguistic stimuli used in the Thai-versus-English-speaker studies were actual words in Thai, while they were only possible words in English. Thus, any differential activation or hemispheric advantage could plausibly be a reflection of some aspect of lexical access, rather than a reflection of tone processing. In fact, GWH raise this possibility in an effort to explain the surprising (from their perspective) result of the tone-minus-baseline subtraction for English speakers: this subtraction did reveal some left frontal activation. GWH attribute this activation to the presence of consonantal and vowel information in the Thai stimuli, which may have been subject to linguistic processing by the English speakers. (Note, however, that this explanation turns the absence of left frontal activation in the tone-minus pitch subtraction into a puzzling mystery, one which GWH ignore.) The same concern confounds the dichotic listening paradigm and the aphasia studies, where the tasks involved lexical decision.

In addition to the potential confound between tone processing and lexical access, results from other types of experiments cast doubt on the validity of the functional theory of hemispheric specialization. First, a lack of right-ear-advantage effects for steady-state vowels in dichotic listening experiments (Shankweiler and Studdert-Kennedy 1967, cited in Van Lancker and Fromkin 1973) shows that the functional theory does not extend to all types of linguistic stimuli; and second, recent brain imaging experiments provide evidence in favor of the acoustic theory of hemispheric specialization, at least with respect to the processing of acoustic transients. For example, Johnsrude et al. 1997 shows, using PET, that the left hemisphere is implicated in the processing of acoustic transient information whether the information is linguistic (e.g., consonantal) or not (e.g., pure tones with frequency glides). Conversely, Gage et al. 1998 shows that the M100 response (as measured by MEG) is sensitive to acoustic differences in the onsets of speech sounds. Comparing stop consonant onsets to non-stop consonant onsets, Gage et al. (1998) found an asymmetry in M100 latencies in the left versus right hemispheres, conditioned by the two different types of onsets: the M100 to stop onsets was faster in the right hemisphere than in the left, while the M100 to non-stop onsets was faster in the left hemisphere than in the right. Gage et al. (1998) attribute this effect to differences in the energy amplitudes of stop versus non-stop onsets, further supporting the hypothesis that differential processing is based on acoustic rather than functional differences.

Moreover, recent hypotheses about the processing of information of different frequencies give us reason to think that low-frequency pitch information might be processed differently from high-frequency vowel and consonant information such as rapid formant transitions (e.g., Johnsrude et al. 1997, Belin et al. 1998, Ivry and Lebby

1998). These studies indicate a right-hemisphere computational bias for processing low-frequency information, and a left-hemisphere bias toward high frequencies.

If it is correct that this bias exists, and if the evidence for acoustic rather than (or in addition to) functional processing is correct, the left-hemisphere advantage in tone languages found in the studies cited must be due to some other factor, such as the potential confound of lexical access mentioned above.

3. MEG and the M100

MEG, which records the magnetic fields generated by electrical activity in the brain, provides the temporal resolution necessary to avoid the confound of lexical access in auditory processing. In contrast to other brain imaging techniques (e.g., PET and functional magnetic resonance imaging (fMRI)), which are limited to a temporal resolution on the order of seconds, MEG is able to provide a millisecond-by-millisecond picture of brain activity (see Roberts and Poeppel 1996 and references cited therein). Studies of lexical access have converged on the conclusion that lexical access is indicated by brain activity peaking at around 350 ms post-onset (Pylkkänen et al. 2000); therefore, we can circumvent the lexical access confound by using MEG's fine temporal resolution to focus on earlier brain activity that is known to indicate auditory processing.

The auditory evoked response peaking approximately 100 ms post-onset, known as the M100, is the perfect object of study for the issues being investigated. First, it is an auditory-evoked response, being localizable to auditory cortex (Kuriki and Murase 1989, Diesch et al. 1996, Poeppel et al. 1997, and others). Second, it is automatic and early, far earlier than any activity implicated in lexical access. Third, the latency and amplitude of the M100 response are known to be sensitive to both the acoustic attributes and the perceptual characteristics of auditory stimuli.

3.1 M100 Latency

The latency of the M100 response to pure tones has been shown to depend upon stimulus frequency (Roberts and Poeppel 1996), with the shortest latency occurring in response to frequencies at around 1000 Hz. For vowels, complex stimuli with energy at different frequencies, the M100 latency appears to be sensitive to the frequency of the first formant (F1), to the exclusion of the fundamental frequency (F0), as has been shown by Poeppel et al. 1997. This study compared the M100 response to a set of vowels presented at both 100 Hz and 200 Hz fundamental frequencies. While the difference in F0 had no significant effect on the M100 latency, differences in F1 did play a role: the M100 latency for /a/ (higher F1) was significantly shorter than the latencies for /i/ or /u/ (lower F1s; see also Diesch et al. 1996). Furthermore, we can infer from studies like Tiitinen et al. 1999, in which the M100 latency to vowels was shown to be significantly longer than the M100 latency to a pure tone whose frequency closely approximated the vowels' F1s, that stimulus complexity also affects M100 latency. Finally, M100 latency may depend upon hemisphere. A number of studies have found that right hemisphere latencies are shorter than left hemisphere latencies, though the difference only approaches significance (e.g., Poeppel et al. 1996 reporting on CV stimuli, and Poeppel et al. 1997 reporting on vowel stimuli).

The studies cited above all examined the M100 response in speakers of non-tone languages, for whom F0 does not carry lexically contrastive information. In tone languages, however, F0 plays a much more central role in the linguistic system. One might therefore hypothesize that there would be an effect of F0 on M100 latencies for speakers of tone languages in response to tone-bearing linguistic stimuli.

3.2 M100 Amplitude

With respect to the amplitude of the M100 response, increases in M100 amplitude have been shown to correspond to increases in stimulus intensity (e.g., Bak, Lebech, and Saermark 1985). Increases in M100 amplitude also result from attention (as required by the experimental task; Poeppel et al. 1996). More relevant to this study is evidence which suggests that M100 amplitudes in the two hemispheres are affected differently by differing types of stimuli—in particular, linguistic versus non-linguistic stimuli. In an experiment comparing German speakers' M100 responses to German vowels versus a 1000 Hz tone, Eulitz et al. (1995) report a weak ($p<0.06$) hemisphere by stimulus type interaction: M100 amplitudes were higher for vowels than tones in the left hemisphere, but lower for vowels in the right hemisphere. (It should be borne in mind, however, that in light of the weakness of this effect, Eulitz et al. 1995 emphasize the inconclusiveness of the result, making sense of it by referring to dichotic listening studies which suggest that vowel processing is not strongly lateralized. See also Poeppel et al. 1996, where it is reported that most studies of M100 amplitude have failed to show any hemispheric asymmetries in response to vowel stimuli.)

The cited studies again only examined the M100 response in speakers of non-tone languages. The present study therefore addresses the question, will speakers of tone languages show differential hemispheric activation, as measured by M100 amplitude? If so, and by examining how their responses differ from speakers of non-tone languages, we might also be able to learn more about the factors that determine the amplitude of the M100 response.

3.3 Using the M100 to Test the Two Theories

The experiment we use to test the functional and acoustic theories of tone processing (described more fully below) involves Chinese and English speakers hearing two types of stimuli: syllables with Chinese tone contours (F0), and isolated F0 contours without any vowel information. The former constitute linguistic stimuli for both groups, but only for the Chinese speakers will the F0 information be linguistically relevant (only in Chinese are the particular F0 contours discrete linguistic units). The isolated tones should be non-linguistic for both populations. Given what we know about the latency and amplitude of the M100 response, the two theories make the following predictions regarding the processing of these stimuli:

Syllables	Latency of M100	Amplitude of M100
Acoustic Theory	NO DIFFERENCE between Chinese and English speakers' M100 latency/ amplitude	
Functional Theory	DIFFERENCE: if Chinese speakers' M100 is sensitive to F0, expect longer LH latency (or greater L-R asymmetry) when processing linguistic tone	DIFFERENCE: assuming higher amplitude = more work, expect Chinese speakers to have higher LH amplitude when processing linguistic tone
Isolated tones	Out of a linguistic context, *both theories* predict that isolated F0 contours will produce the same responses in both Chinese and English speakers.	

4. Materials and Methods

For the present experiment two types of auditory stimuli were constructed: syllables and tones. The syllables were three naturally produced VC syllables recorded by a native speaker of Mandarin Chinese. The three were identical in segmental content and differed only in the tone that they carried. The phonotactically possible Mandarin word *ang* was recorded with tone 1 (level high tone, H), tone 2 (rising tone, LH), and tone 4 (falling tone, HL). Only with tone 2 is *ang* a real word in Mandarin (meaning 'to raise'); with other tones it only occurs in combination with other syllables. In order to match the syllables as closely as possible for total duration, the speaker repeated each target syllable several times to a constant rhythm determined by a metronome. (Tone 3, falling-rising, was not used because syllables with Tone 3 are generally longer.) The syllables were digitized using Speech Station (Sensimetrics). Stimulus duration was measured for each token, and one token from each type was selected for inclusion in the stimulus set. The tokens selected were those closest in duration to 300 ms with the clearest F0 component (i.e., with the least amount of F0 distortion due to glottal pulse "creakiness"). Tokens were clipped to 300 ms starting at the beginning of steady F0 information. The first and last 10 ms were faded in and out respectively using SoundEdit16. The edited stimuli were correctly identified and discriminated by a fluent Chinese speaker when played in isolation (and by the Chinese subjects in the experiment), thus ensuring that the stimuli were good exemplars of the three target tones.

Based on the tone information carried on the fundamental frequencies of the syllable stimuli, three synthetic tones were created, each of 300 ms duration, which exactly mimic the F0 of the corresponding syllable stimuli. Again, the first 10 ms of each stimulus were faded in. The fade-out portion at the end of each stimulus varied slightly: it was equal in duration to the interval between the end of steady F0 information in the corresponding syllable stimulus and the 300 ms point (approximately 40 ms for each stimulus). The synthesized tones were also correctly identified and discriminated by the same fluent Chinese speaker when played in isolation.

Figure 1: F0 Contours

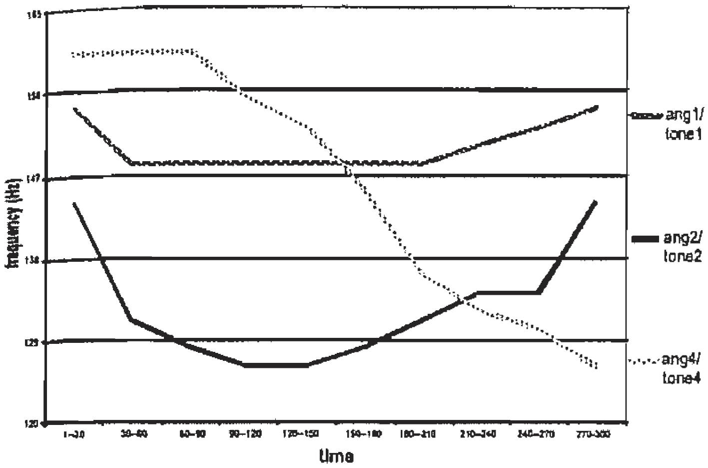


Figure 1 plots F0 for both the syllables and corresponding tones. The F0 of *ang1* and tone1, high level tone, remains relatively stable throughout; it begins and ends at 152 Hz, with a steady 148 Hz middle portion. *Ang2* and tone2, rising tone, start lower than the H tone, at 144 Hz, dip to a steady-state portion at 127 Hz, and rise back up to 141 Hz at the end of the syllable. *Ang4* and tone4 begin at the highest frequency, 160 Hz, and start to fall gradually after approximately 75 ms until a low of 126 Hz at the end of the syllable.

Twelve paid volunteers, six native English speakers (one female) and six native Mandarin Chinese speakers (four females), all students and employees at MIT, all of whom gave informed consent, participated in the experiment. (Three additional Chinese speakers also took part but were excluded due to problems with the data.) All subjects were right-handed and had no history of hearing or neurological disorders.

Subjects lay supine in a magnetically shielded room while stimuli were presented binaurally over earphones. Evoked magnetic fields were recorded using the MIT/KIT 64-sensor whole-head biomagnetometer array. Recording was continuous at a sampling rate of 500 Hz.

Stimuli were presented in two experimental conditions (order randomized across subjects): an ID condition and a passive listening condition. Before beginning these experimental conditions, subjects listened to a 1 kHz tone presented 100 times (for the purpose of identifying and localizing the M100, as well as providing a comparison for the syllables and tones). In the ID condition (broken into two blocks), subjects were instructed to press one button when they heard that same 1 kHz tone, and a second button if they heard any other sound. In each of the two blocks, 51 instances of the same 1 kHz tone were presented randomly along with 51 tokens of each of the three syllable and three tone stimuli. Between blocks the subject was given a short break. In the passive condition, subjects listened to two blocks of 51 instances of each of the syllable and tone stimuli, again with a break between blocks. Each stimulus was therefore presented a total of 102 times in each condition.

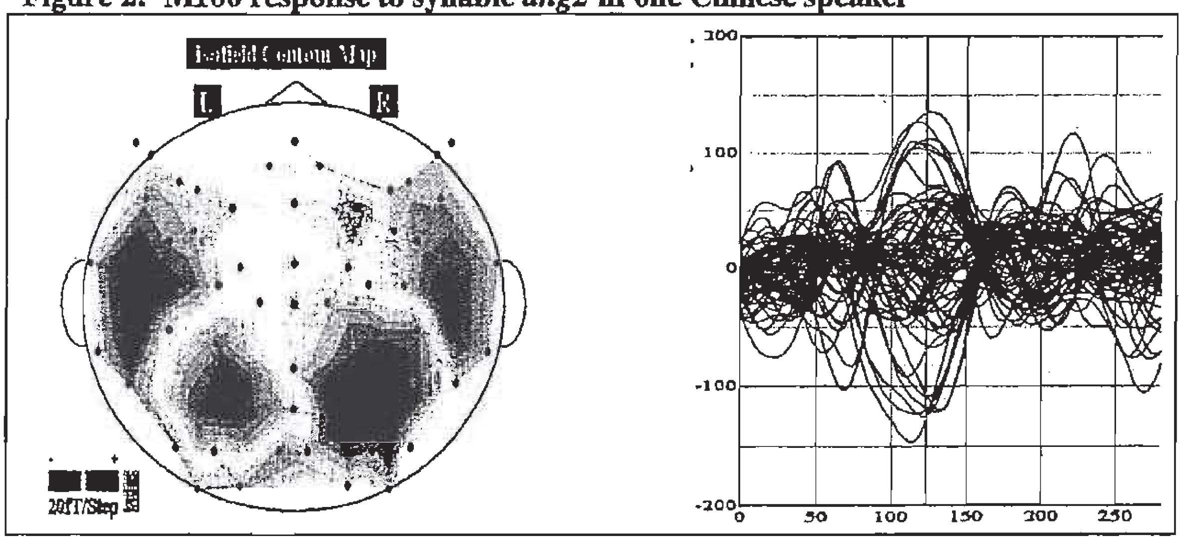
The ID condition was meant to test the effect of attention and task demands on the M100. However, it turned out to be very difficult to locate the M100 to many of the stimuli in the ID condition for several of the subjects, and in addition movement from the button presses caused some interference in the data. Therefore only the passive listening condition is reported here.

The raw data were subjected to a noise reduction routine to eliminate known electrical activity from external sources. The data for each of the stimuli were then averaged separately, in 600 ms windows keyed to the onset of the stimulus: 100 ms pre-, 500 ms post-onset. The averaged signal was subjected to a low-pass filter and adjusted to baseline using a 100 ms pre-stimulus interval.

5. Results

All subjects evidenced a clear M100 response to most of the stimuli. Figure 2 shows the M100 response to the syllable *ang2* for one subject (Chinese speaker). The graph on the right overlays recordings from each sensor (each represented by a line) across time. A large peak occurs in the interval between 100 and 150 ms. A spatial layout on the left provides the sensor readings (in the form of a contour map) for a single moment in time, that of the vertical line between 100 and 150 ms in the graph on the right. Two dipolar patterns are easily discernible, centered over left and right auditory cortex (classic M100 response). The response to the 1 kHz tone was often most robust, and was used to select channels utilized in computing the root mean square (RMS) of the magnetic field strength. 16 channels in each hemisphere were selected by visual inspection as those centered over the area of greatest activity. The M100 was taken to be the RMS peak over these channels in the latency range 80–170 ms.

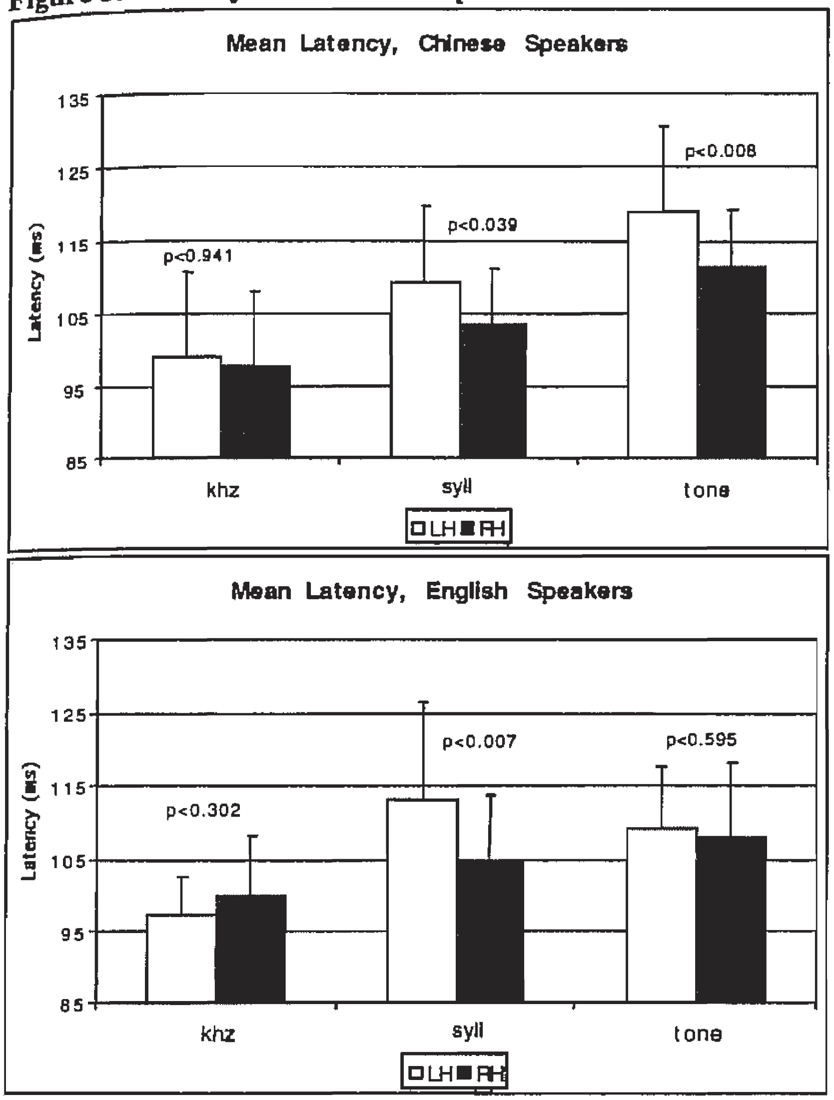
Figure 2: M100 response to syllable *ang2* in one Chinese speaker



5.1 Latency

Figure 3 plots the latency of the M100 response by language background (Chinese vs. English), stimulus type (syllable vs. tone, plus the 1kHz tone), and hemisphere.

Figure 3: Latency of the M100 response



Statistical analysis (ANOVA, factors: language background, stimulus type, hemisphere) indicates a main effect of hemisphere for both Chinese and English subjects: the RH was significantly faster than the LH ($p<0.005$). There was also a language x stimulus type interaction ($p<0.045$) and a hemisphere x language x stimulus type interaction that approached significance ($p<0.078$). Post-hoc Scheffe tests show that both the Chinese and English groups were faster in the RH than in the LH in response to syllables; however, while the Chinese group was significantly faster in the RH than in the LH in response to tones ($p<0.0074$), the English group showed no such RH versus LH asymmetry. In other words, the Chinese speakers showed the same asymmetry to tones that they did to syllables, but the English speakers had no asymmetry to tones.

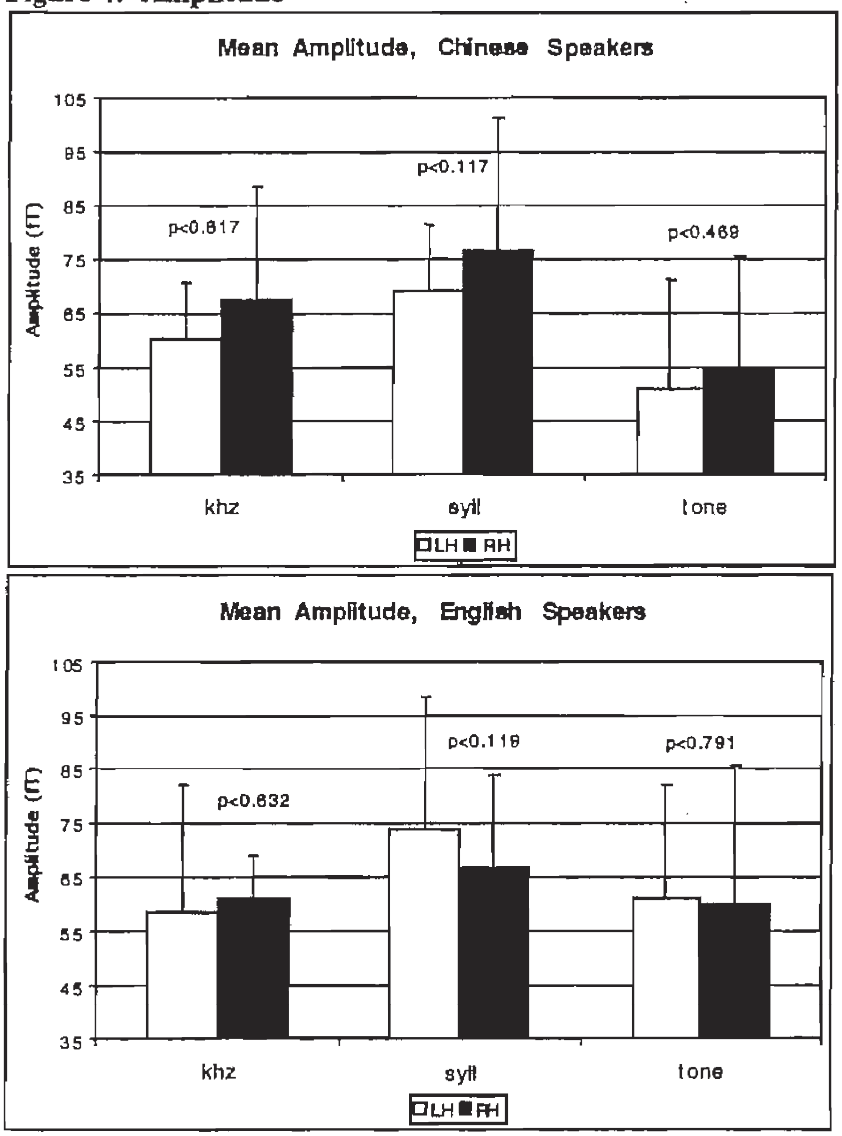
These latency findings disconfirm both the acoustic and the functional theories. The acoustic theory predicts that both language groups should have the same response patterns to both syllables and tones. As the interactions reveal, however, the groups differ: the Chinese speakers showed the same asymmetry to both stimulus types, but the

English speakers differed in their responses to the different stimulus types. As for the functional theory, it predicts that the two language groups should have the same response patterns to tones, but different responses to syllables. This is also false: the two groups patterned similarly on the syllable condition, but diverged on the tone condition.

5.2 Amplitude

Figure 4 plots the amplitude of the M100 response by language background (Chinese vs. English), stimulus type (syllable vs. tone, plus the 1kHz tone), and hemisphere.

Figure 4: Amplitude



Statistical analysis (ANOVA, factors: hemisphere, language, stimulus type) revealed a significant difference in the M100 amplitude as a function of stimulus type: the syllables elicited higher-amplitude M100s than did the tones ($p<0.003$). (Because we did not control for stimulus intensity, it is difficult to draw any conclusions from this finding.) In addition, there was a weak hemisphere x language x stimulus type (syllable vs. tone) interaction ($p<0.263$). For both populations, the RH and LH M100s were roughly equal in amplitude in response to tone stimuli. In response to syllables, however, the Chinese and English groups showed a reversal in RH versus LH amplitude asymmetries. The Chinese RH response to syllables was greater in amplitude than the LH response (by 7.2 fT), though this difference only approached significance ($p<0.117$). In contrast, the English LH evidenced a higher-amplitude response than did the RH (again by 7.2 fT), though again this difference only approached significance ($p<0.119$). The Chinese RH syllable response was also significantly higher than the English RH response ($p<0.041$), even though the Chinese M100 was lower in amplitude than the English in all other comparisons.

These results falsify the acoustic theory, which predicted that both language groups should have the same response patterns to both syllables and tones. We found instead that the two groups had similar responses to tones, but showed hemispheric differences in *opposite directions* in response to syllables. The functional theory predicted this finding: that both language groups should have the same response patterns to tones, but different responses to syllables.

5.3 Interpreting the Results

The major points of interest can be framed as follows: the amplitude results for the Chinese and English speakers in the syllable condition contrast with one another, and the latency results for the English speakers in the tone condition contrast with the pattern found in the other latency results. We speculate that the first observation can be interpreted as a specific effect of native language background on linguistic processing, and that the second observation can be interpreted as a difference in auditory processing of linguistic versus non-linguistic acoustic stimuli.

Beginning with the amplitude figures, the results imply that the English speakers' LH is dominant in the early stages of processing in the syllable condition, while the Chinese speakers' RH is dominant. Roughly, we consider amplitude to relate to the amount of "work" being carried out, and understand hemispheric dominance as (near) statistically significant higher amplitude in one hemisphere. A straightforward approach to the reversal in hemispheric dominance in the syllable condition is to assume that the difference is due to language background. This is not a contentious assumption since the stimuli presented to the two groups were identical. It is something about acquiring Chinese as a native language, then, that accounts for the Chinese speakers' RH dominance at the early stages of linguistic processing, and something about acquiring English as a native language that accounts for the English speakers' LH dominance at the early stages of linguistic processing. What might that be? We suggest that it is the fact that Chinese is a tone language, while English is not. The Chinese speaker must come to treat the lower frequencies of the linguistic acoustic signal quite differently from the English speaker. We suggest that the importance of low-frequency F0 information as tonal information is reflected by higher early RH involvement, or dominance, in the tone language speaker. In non-tone language speakers, such as our English subjects, where lexical and phonological contrasts in the acoustic signal are typically high-frequency segmental information, we find LH dominance.

These patterns seem to fit in with the generalizations about hemispheric biases for information of different frequencies, discussed above with regard to the acoustic theory, but the connection is obviously tenuous at this point. We must, then, continue to confine our speculation about the exigencies of linguistic tone processing on the RH to the very earliest stages of auditory processing. In effect, we find that linguistic tone processing, at least at the very early stages measured by the M100, is not LH dominant. In this sense, the acoustic theory is borne out, and the functional theory is falsified, in that the Chinese speakers' need to take account of F0 at a lower linguistic level begins very early with high RH involvement in auditory processing, and not with high LH involvement just because tone information is linguistic. The RH of the non-tone language speaker is not similarly crucially involved at early stages. Interestingly, rather than a lack of hemispheric dominance, which might be the most conservative prediction for the English speakers' M100 amplitude given the above explanation for Chinese speakers' RH dominance, we actually find LH dominance. We suggest that just as the processing demands of the tone language early on place relatively greater importance on low-frequency information, a RH specialty on the acoustic theory, the processing demands of the non-tone language early on place relatively greater importance on high-frequency information, a LH specialty according to the acoustic theory.

Moving on to the latency results, our finding for both the Chinese and the English speakers in the syllable condition is that the M100 peak latency is significantly earlier in the RH than the LH. Note that to the subjects of both groups, the syllables sound like language; the English subjects may not easily identify or discriminate the Mandarin tones, but the VC syllables are nonetheless obviously linguistic stimuli. By contrast, in the tone condition only the Chinese speakers showed the LH-RH asymmetry with significantly faster RH peak latency. The English speakers showed no hemispheric latency asymmetry in the tone condition. Put another way, the Chinese speakers treated the pure tone stimuli of the tone condition just like they and the English speakers treated the linguistic stimuli of the syllable condition—their M100 latency evidenced the LH-RH asymmetry familiar from the syllable condition (and reported many places in the M100 literature as a general tendency). Why should the latency asymmetry disappear for the English speakers in the tone condition? That is, what is it about the tone stimuli for the English speakers that can explain the absence of the LH-RH latency asymmetry?

One explanation is that the hemispheric asymmetry is a characteristic of linguistic processing. Its absence in the tone condition for the English speakers means that they do not treat the tone stimuli as linguistic, whereas the tone language speakers do. On this approach, the Chinese speakers treat the pure tone stimuli similarly to the linguistic stimuli of the syllable condition, but the English speakers do not. In this sense, we can understand the LH-RH asymmetry to index linguistic auditory processing; in non-linguistic auditory processing, the RH M100 latency is not significantly earlier than the LH.

One initial source of supporting evidence for this position comes from our results on the 1kHz pure tone stimuli, which showed no hemispheric asymmetry in either subject group (Figure 3). However, M100 latencies are at their fastest at frequencies this high, so it is possible that the absence of a hemispheric asymmetry in the 1kHz tone stimuli results from a sort of ceiling effect. For further confirmation of the conjecture that a faster RH M100 peak latency is a characteristic of linguistic auditory processing, we must test the prediction that Chinese speakers' LH-RH latency asymmetry will disappear in a pure tone condition with stimuli lower than 1kHz but higher than the human F0 range. In this case, the pure tone stimuli would neither be a possible discrete linguistic unit, nor at a frequency where hemispheric differences would be neutralized. Alternatively, one could present Chinese subjects with tones created by musical instruments; even at frequencies which carry tonal information in Mandarin, the stimuli would be recognizably non-linguistic, and according to the above speculations, should not show a LH-RH latency asymmetry (Martha McGinnis, p.c.).

If it is true that linguistic stimuli demonstrate a LH-RH latency asymmetry, but non-linguistic stimuli do not, then there is a crucial sense in which the functional, or speech-is-special, view must be entertained. We have already seen in relation to M100 amplitude that this is not true for the early stages of processing, in which the Chinese subjects showed RH dominance, but it does seem to be true in terms of the latency of the M100 peak. It is very difficult to say what the M100 peak latency corresponds to in terms of speech processing, and, as such, it is also difficult to explain why a hemispheric latency asymmetry should be expected for linguistic auditory stimuli alone, but we will offer a possible approach at this point. It could be that language processing requires a coordination of the acoustic information being processed by the two hemispheres. That is, if the two hemispheres process different frequencies of the acoustic signal, then, if the information carried on these frequencies is to be interpreted in a unified way, as language, it must be (re)assembled for higher-level processing. Since the RH processes slow-changing, lower frequency information, it must sample at longer cycles than the LH, which processes faster-changing, high-frequency information. While this might suggest just the opposite of the pattern we see, that is, a faster LH M100 peak, there is actually good reason to expect the asymmetry we find. There is voluminous evidence that the LH is specialized for language, and we have no reason to doubt that the LH language areas are dominant for higher-level linguistic processing. There is a sense, then, in which the LH must wait to receive any linguistic information that is processed in the RH before higher-level computation can proceed. If crucial aspects of a language are carried on acoustic frequencies for which the RH is responsible, then the computations of the LH language areas must be delayed until that information can be conveyed to the LH. Thus, neural responses that are correlated with auditory processing, such as the M100, will be delayed in the LH until their RH analogues are available. In non-linguistic processing, when there is no need to assimilate acoustic information from the RH before higher level LH computations, the LH auditory responses need not be delayed. Though this theory is entirely speculative, it does offer an approach to understanding the very interesting pattern of hemispheric (a)symmetry that emerged in this study.

6. Conclusion

We did find an effect of language background, contradicting the acoustic theory, but it was not precisely the effect predicted by the functional theory. The latency results suggest that speech is processed differently from non-linguistic stimuli, and the amplitude results suggest that linguistic tone processing involves the RH. A refinement of the functional theory that incorporates a crucial insight of the acoustic theory is thus called for: the functional theory is correct, in that speech is special. When a stimulus is treated as linguistic (which is determined by language background), it will elicit a LH-RH M100 latency asymmetry. However, the acoustic properties of a linguistic stimulus determine differential hemispheric involvement, or dominance. In particular, low frequencies show higher RH M100 amplitude.

References

Bak, C. K., J. Lebech, and K. Saermark (1985). Dependence of the auditory evoked magnetic field (100 msec signal) of the human brain on the intensity of the stimulus. *Electroencephalography and Clinical Neurophysiology* 61: 141-149.

Belin, Pascal, Monica Zilbovicius, Sophie Crozier, Lionel Thivard, Anne Fontaine, Marie-Cecile Masure, and Yves Samson (1998). Lateralization of speech and auditory temporal processing. *Journal of Cognitive Neuroscience* 10: 536-540.

Diesch, Eugen, Carsten Eulitz, Scott Hampson, and Bernhard Ross (1996). The neurotopography of vowels as mirrored by evoked magnetic feld measurements. *Brain and Language* 53: 143-168.

Eulitz, Carsten, Eugen Diesch, C. Pantev, Scott Hampson, and T. Elbert (1995). Magnetic and electric brain activity evoked by the processing of tone and vowel stimuli. *Journal of Neuroscience* 15: 2748-2755.

Gage, Nicole, David Poeppel, Timothy P. L. Roberts, and Gregory Hickok (1998). Auditory evoked M100 reflects onset acoustics of speech sounds. *Brain Research* 814: 236-239.

Gandour, Jack and Rochana Dardarananda (1983). Identification of tonal contrasts in Thai aphasic patients. *Brain and Language* 18: 98-114.

Gandour, Jack, Donald Wong, and Gary Hutchins (1998). Pitch processing in the human brain is influenced by language experience. *NeuroReport* 9: 2115-2119.

Gandour, Jack, Donal Wong, Li Hsieh, Bret Weinzapfel, Diana Van Lancker, and Gary D. Hutchins (2000). A crosslinguistic PET study of tone perception. *Journal of Cognitive Neuroscience* 12: 207-222.

Ivry, Richard and Paul C. Lebby (1998). The neurology of consonant perception: Specialized module or distributed processors? In Mark Beeman and Christine Chiarello (eds.), *Right Hemisphere Language Comprehension: Perspectives from Cognitive Neuroscience* , pp. 3-25, Mahwah, New Jersey: Erlbaum.

Johnsrude, Ingrid S., Robert J. Zatorre, Brenda A. Milner, and Alan C. Evans (1997). Left-hemisphere specialization for the processing of acoustic transients. *NeuroReport* 8: 1761-1765.

Kuriki, S. and M. Murase (1989). Neuromagnetic study of the auditory responses in right and left hemispheres of the human brain evoked by pure tones and speech sounds. *Experimental Brain Research* 77: 127-134.

Packard, Jerome L. (1986). Tone production deficits in nonfluent aphasic Chinese speech. *Brain and Language* 29: 212-223.

Poeppel, David, Colin Phillips, Elron Yellin, Howard A. Rowley, Timothy P. L. Roberts, and Alec Marantz (1997). Processing of vowels in supratemporal auditory cortex. *Neuroscience Letters* 221: 145-148.

Poeppel, David, Elron Yellin, Colin Phillips, Timothy P. L. Roberts, Howard A. Rowley, Kenneth Wexler, and Alec Marantz (1996). Task-induced asymmetry of the auditory evoked M100 neuromagnetic field elicited by speech sounds. *Cognitive Brain Research* 4: 231-242.

Pylkkänen, Liina, Elissa Flagg, Andy Stringfellow, and Alec Marantz (2000). A neural response sensitive to priming: An meg study of lexical access. Poster, BioMag2000, Helsinki University of Technology, Finland.

Roberts, Timothy P. L. and David Poeppel (1996). Latency of auditory evoked M100 as a function of tone frequency. *NeuroReport* 7: 1138-1140.

Shankweiler, D. and M. Studdert-Kennedy (1967). Identification of consonants and vowels presented to left and right ears. *Quarterly Journal of Experimental Psychology* 19: 59-63.

Tiitinen, Hannu, Päivi Sivonen, Paavo Alku, Juha Virtanen, and Risto Näätanen (1999). Electromagnetic recordings reveal latency differences in speech and tone processing in humans. *Cognitive Brain Research* 8: 355-363.

Van Lancker, Diana and Victoria A. Fromkin (1973). Hemispheric specialization for pitch and "tone": Evidence from Thai. *Journal of Phonetics* 1: 101-109.

Van Lancker, Diana and Victoria A. Fromkin (1978). Cerebral dominance for pitch contrasts in tone language speakers and in musically untrained and trained English speakers. *Journal of Phonetics* 6: 19-23.

Department of Linguistics
MIT E39-245
77 Massachusetts Avenue
Cambridge, MA 02139

muawiya@mit.edu
eflagg@mit.edu
vilin@mit.edu